



Samuel Oyerinde

 Lagos, Nigeria  oyerindesamuelabiodun@gmail.com  08064427860  Github Profile

 Google scholar  Huggingface

PROFESSIONAL SUMMARY

Applied ML engineer with expertise in speech technologies for low-resource African languages, including Text-to-Speech (TTS) and Automatic Speech Recognition (ASR). Experienced in multilingual representation learning, audio data pipelines, and dataset curation, with a strong interest in deploying NLP solutions for social good, language inclusion, and cultural preservation.

EDUCATION

Computer Science (BSc), Federal University of Agriculture, Abeokuta Apr 2018 – Mar 2023
Abeokuta, Nigeria

WORK EXPERIENCE

NLP & Audio Data Processing, Masakhane Research Foundation (DDD) Apr 2025
Contract, remote

- Coordinated data workflows across multiple teams, improving delivery speed and documentation accuracy.
- Designed structured tracking systems for progress monitoring and reporting.
- Collaborated with cross-functional teams (linguists, engineers, data processors) to ensure timely project delivery.
- Machine Learning Support: Utilize Hugging Face Transformers for data processing and NLP model training.
- Led technical oversight of data quality assurance processes, ensuring accuracy and consistency across multiple language datasets.
- Managed system operations for audio recording infrastructure and data processing workflows on Windows and MacBook environments.
- Performed High-level data analysis and visualisation with Pandas, Seaborn, and Matplotlib to make core business decisions.

Audio Engineer, DSNai Jan 2025 – Apr 2025
Lagos, Nigeria

- Audio Processing This involves reducing noise, balancing frequencies, and fine-tuning levels, achieving a 60% improvement in sound clarity for speech model training .
- Managed large-scale audio projects and ensured compliance with technical and program objectives.
- Documented processes, tracked deliverables, and reported metrics to project leads.
- Adapted to evolving project needs, proactively supporting additional tasks to align with organizational goals.

Frontend Engineer, Optimus AI Labs May 2023 – Sep 2024
Lagos, Nigeria

- Supported frontend team coordination, documentation, and iterative development reviews.
- Cross-Browser Compatibility: Delivered a consistent experience across devices by ensuring seamless cross-browser compatibility using CSS media queries, leading to a 20% reduction in user-reported issues.
- Tools: Reactjs, SWR, Javascript, DevTools, Formik, Yup, MUI, Tailwind, SwiperJs, Typescript, Recharts.

ML Engineer, Freelance Oct 2021
Lagos, Nigeria

- **Feature Engineering:** Grouping and merging categories to simplify classification tasks.

- **Trained multilingual NLP models** using PyTorch and HuggingFace Transformers for text classification tasks in low-resource African languages.
- **Collaborations:** *Collaborated with cross-functional teams to deploy models.*

AI Instructor, Risenetwork

Mar 2021 – May 2021
Mowe, Ogun State.,
Nigeria

- Trained and mentored over 300 young graduates at RCCG Youth Camp, teaching them how to leverage Python for artificial intelligence applications.
- Designed and delivered engaging lessons, ensuring effective and efficient learning while providing one-on-one support.
- Guided students in developing a project on predicting customer churn, resulting in the best presentation for the entire program.
- Areas: Data Visualization, Feature Engineering, Modeling, and Prediction.

Campus Ambassador, FUNAAB, Grazac

Feb 2021 – Feb 2022
Abeokuta, Nigeria

SELECTED COMMUNITY SERVICES

AI+ FUNAAB Core Team Member, Data Science Nigeria

2019 – Mar 2022
Abeokuta, Nigeria

- Collaborated with fellow students to train and mentor peers in data science and machine learning.
- Contributed to the growth and development of the AI+ FUNAAB community, fostering a culture of innovation and collaboration.

Programming Instructor, Ogun Innovate

2020 – 2021
Abeokuta, Nigeria

- Trained and tutored over 30 students in coding skills, including Python and Javascript, while ensuring a productive and engaging learning environment.
- Provided individualized support and feedback to students, fostering a supportive and inclusive learning community.

Program Coordination & Administration

- **Program & Coordination Tools:** Google Workspace, Airtable, Slack, Trello, ClickUp
- **Documentation & Reporting:** Excel (advanced formulas), progress tracking, report writing
- **Communication:** Virtual coordination, remote collaboration, stakeholder follow-ups

PUBLICATIONS & COLLABORATIONS

The African Stopwords project: curating stopwords for African languages, Chris Emezue, Hellina Nigatu, Cynthia Thinwa, Helper Zhou, Shamsuddeen Muhammad, Samuel Oyerinde et al. 2023

AfriWOZ: Corpus for Exploiting Cross-Lingual Transferability for Generation of Dialogues in Low-Resource, African Languages, Tosin Adewumi, Mofetoluwa Adeyemi, Oyerinde Samuel et al. 2022

MasakhaNER: Named Entity Recognition for African Languages, David Ifeoluwa Adelani, Jade Abbott, Graham Neubig, Oyerinde Samuel and 48 more authors. MasakhaNER: Named Entity Recognition for African Languages . In AfricaNLP Workshop at ACL, 2021 [🔗](#) 2021

SKILLS

Programming

- Python, Javascript, Typescript.

Frameworks

- Scikit-learn, Google Sheets, PyTorch, Hugging Face Transformers.

Certificates

AI BOOTCAMP CERTIFICATE 2020 [↗](#) (Data Science Nigeria) •

Responsive Web Design [↗](#) (FreeCodeCamp) • **Generative AI and LLMs: Architecture and Data Preparation** [↗](#)

RESEARCH EXPERIENCE

Masakhane Research – The African Stopwords Project - Curating Stopwords for African Languages | Curated Pidgin stopwords for African languages, contributing to the development of a comprehensive stopwords resource | Ensured the quality and relevance of data used for training models, enhancing the performance of AI systems | Advanced NLP research for African languages, enabling more accurate text processing and analysis | 2023

Masakhane Research – AfriWOZ - Corpus for Exploiting Cross-Lingual Transferability for Generation of Dialogues in Low-Resource African Languages | Trained and performed dataset curation for the seq2seq model used in this research project | Contributed to the development of a corpus for African languages, enabling cross-lingual transferability and dialogue generation | 2022

Masakhane Research – Named Entity Recognition for African Languages | Collaborated on the largest language annotation effort for Nigerian Pidgin, with over 75 million speakers in Africa | Contributed to the development of NLP models for African languages, addressing underrepresentation in NLP research | Annotated datasets for training and evaluating NLP models, ensuring high-quality data for improved model accuracy | 2021

Programming – *3+ years of Python programming for ML projects, including data preprocessing, model training, and evaluation.*

Languages

English, Yoruba, Nigeria-Pidgin